






Bridge Structural Damage Identification Using Causal-Invariant Spatio-Temporal Representation Learning

Shihong Huang¹ , Shenghuan Qin^{1*} , Chengye Liang² 

¹ Department of Management Science and Engineering, Guangxi University of Finance and Economics, Nanning 530000, China

² Department of Architectural Engineering, Guangxi Electrical Polytechnic Institute, Nanning 530000, China

Corresponding Author Email: qinsh@gxufe.edu.cn

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420648>

ABSTRACT

Received: 18 April 2025

Revised: 31 October 2025

Accepted: 17 November 2025

Available online: 31 December 2025

Keywords:

Structural health monitoring; causal invariance; spatio-temporal representation learning; graph neural networks; vibration signal analysis

Bridge structural health monitoring (SHM) based on vibration signals is strongly affected by variations in operating conditions, environmental disturbances, and limited labeled data. Under such circumstances, data-driven models tend to exploit condition-dependent correlations rather than damage-related mechanisms, which leads to unstable performance and limited generalization across different operating scenarios. To address this problem, a causal-invariant spatio-temporal representation learning (CISRL) framework is developed for bridge damage identification and localization. The framework integrates three components within a unified architecture: spatio-temporal feature extraction from multi-sensor vibration signals, graph-based modeling of structural damage propagation along the bridge topology, and cross-condition invariance regularization to suppress condition-specific features. The invariance constraint guides the learning process toward representations that remain stable across different operating conditions while preserving sensitivity to structural damage. The proposed method is evaluated on the Japanese Old ADA bridge dataset and several public multivariate time-series datasets. The results show consistent improvements over existing deep learning approaches in both damage identification and localization tasks, particularly under cross-condition testing, small-sample training, and sensor layout variation. The findings indicate that incorporating causal invariance into spatio-temporal and graph-based learning provides a reliable and practical approach for robust structural damage identification in complex engineering environments.

1. INTRODUCTION

Bridges, as key infrastructure in modern transportation systems, play a critical role in the efficiency of transportation and public safety. However, during their long-term service, bridge structures are inevitably affected by multiple factors such as repeated vehicle loads, temperature and humidity variations, corrosive environments, and material fatigue aging, which can easily lead to structural damage such as crack propagation, stiffness degradation, loosening, or even fracture of components [1, 2]. These types of damage are often characterized by their hidden nature, slow progression, and severe consequences. If not identified and accurately located in the early stages, they may lead to rapid deterioration of structural performance or even catastrophic accidents. Therefore, developing high-precision, high-robustness, and engineering-applicable bridge SHM methods has always been an important research topic in the fields of structural engineering and mechanical systems [3].

Among various SHM technologies, vibration signal-based damage identification methods have received widespread attention due to their ability to reflect the overall dynamic characteristics of the structure, their applicability to complex

structural systems, and ease of engineering implementation [4, 5]. The dynamic response of bridge structures changes in terms of frequency distribution, energy transfer paths, and time-series evolution patterns between the healthy and damaged states, providing a physical basis for damage identification based on vibration signals. Traditional research mainly relies on signal processing methods such as modal parameter changes, frequency domain analysis, wavelet transform, and statistical feature extraction, using manually designed damage-sensitive indicators to characterize changes in the structural state [6-8]. However, these methods typically depend on prior experience and specific operating condition assumptions, and their feature stability and generalization ability are often significantly limited when faced with complex, non-stationary, and noisy environments [9].

In recent years, with the improvement of computational capabilities and the development of sensor technologies, deep learning methods have gradually been introduced into the field of SHM. End-to-end methods based on convolutional neural networks, recurrent neural networks, and transformers are able to automatically learn high-level features directly from raw vibration signals, showing performance advantages over traditional methods in damage identification and localization

tasks [10–13]. Especially in multi-sensor scenarios, deep models can integrate time-series information from different measurement points, mining potential spatial correlations and dynamic evolution patterns from the data, providing a new technical path for overall state assessment of complex structures [14].

At the same time, the increasing scale and density of bridge sensor networks have further amplified the complexity of vibration data, making the learning task not only high-dimensional but also highly heterogeneous across different operational and environmental conditions. This has brought new challenges to the stability, interpretability, and transferability of data-driven SHM models, especially when they are deployed in long-term monitoring scenarios where the operating conditions cannot be fully controlled or anticipated.

Nevertheless, most existing deep learning methods are still based on a correlation-driven learning paradigm, i.e., learning the statistical correlation between input signals and damage labels by minimizing prediction errors [15]. In practical engineering environments, bridge vibration signals are not only affected by structural damage but are also strongly dependent on various non-structural factors such as changes in vehicle speed, traffic flow characteristics, environmental noise levels, and sensor configuration. These factors often exhibit strong correlations with damage labels in specific datasets, leading the model to learn "spurious correlation features" that are discriminative but unrelated to the actual damage [16]. When there is a distribution shift between the test and training conditions, these features quickly fail, leading to a significant decline in model performance. This issue has become an important bottleneck limiting the engineering application of deep learning methods [17].

From a causal inference perspective, bridge structural damage is the fundamental cause of changes in dynamic response, while factors such as vehicle speed, environmental noise, and sensor configuration can be regarded as external disturbances or confounding variables [18]. Although these disturbances affect the specific manifestation of the observed signal, the mechanism of damage's effect on the structural response should remain relatively stable under different conditions. Therefore, if the model can focus on these causal features that are invariant across conditions, rather than relying on statistical correlations in specific environments, it is expected to significantly improve generalization ability and reliability in complex real-world scenarios [19].

Based on this understanding, causal learning and invariant representation learning have gradually gained attention in the fields of machine learning and mechanical systems in recent years. By introducing cross-environment consistency constraints during model training, the impact of environmental interference factors can be effectively suppressed, allowing the model to learn more essential and stable representations [20]. However, systematically incorporating causal invariant learning into the field of SHM, especially in conjunction with multi-sensor spatio-temporal modeling and structural damage propagation mechanisms, still requires further research.

In particular, existing studies rarely address, within a single unified framework, the simultaneous requirements of (i) modeling the spatio-temporal dynamics of multi-sensor vibration signals, (ii) capturing the physical propagation of damage effects along the structural topology, and (iii) enforcing representation invariance across different operating conditions. The absence of such an integrated perspective limits both the robustness and the physical interpretability of

current data-driven SHM approaches.

Therefore, this paper proposes a CISRL method for SHM. Based on deep spatio-temporal feature modeling, this method introduces a damage causal propagation graph to explicitly characterize the spatial diffusion mechanism of damage effects in the structure. At the same time, by incorporating cross-condition invariance regularization constraints, the model is guided to focus on stable features directly related to the damage mechanism, thus constructing a bridge damage identification model with stronger generalization ability and engineering applicability.

2. RELATED WORK

In SHM, vibration response-based damage identification has formed a relatively mature technical spectrum, evolving from traditional signal processing and modal analysis to data-driven deep learning and graph learning frameworks. However, in practical engineering, distribution shifts caused by operational conditions (vehicle speed, load levels, environmental noise, traffic flow, etc.) still serve as the core bottleneck limiting the practical implementation of models. This section reviews relevant research progress in vibration damage identification, spatio-temporal deep modeling, GNN and structural topology modeling, cross-domain/cross-condition generalization, and causal invariant learning, in line with the objectives of this paper, namely "cross-condition generalization + causal invariant representation + graph propagation mechanism modeling."

Although these research directions have been extensively studied in isolation, they are rarely integrated within a single framework that simultaneously addresses spatio-temporal dynamics, structural topology, and cross-condition robustness, which motivates the structured review and positioning presented in this section.

2.1 Vibration-based bridge damage identification methods

Traditional bridge damage identification methods are mainly based on variations in modal parameters (such as natural frequencies, mode shapes, and damping ratios) or statistical and spectral features extracted from vibration signals, which are then used to construct damage indicators and combined with threshold-based criteria or shallow classifiers for damage detection. These approaches offer advantages such as clear physical interpretability and relatively low implementation cost; however, they are sensitive to non-stationary behavior, measurement noise, and complex boundary conditions. In addition, their reliance on manually designed features limits their adaptability to real-world scenarios involving varying operating conditions and heterogeneous multi-sensor layouts.

With the increasing deployment of high-density sensor networks and long-term monitoring systems, data-driven methods have gradually become more prevalent in SHM applications. In this context, SHM problems have been reformulated as multivariate time-series classification, retrieval, or anomaly detection tasks, and their feasibility has been demonstrated using real bridge monitoring data. Publicly available datasets with controlled damage scenarios, such as those from the Old ADA steel truss bridge, have played an important role as benchmark platforms for evaluating cross-condition performance and generalization capability [21–23].

2.2 Deep learning approaches in bridge SHM

Deep learning has achieved notable progress in SHM tasks such as damage detection, localization, and structural state identification by automatically learning damage-sensitive representations in an end-to-end manner. Representative approaches can be broadly categorized as follows:

(1) One-dimensional convolutional neural networks (1D-CNNs): These models extract local patterns from vibration time series through convolutional kernels, providing stable training behavior and efficient inference. However, they are limited in capturing long-range temporal dependencies and cross-sensor coupling effects.

(2) Recurrent sequence models (LSTM/GRU): These networks are more effective in modeling long-term temporal dependencies, but they often suffer from optimization difficulties when applied to long sequences and noisy measurement environments.

(3) Time–frequency representations and multi-scale fusion: In this paradigm, vibration signals are first transformed into the time–frequency domain using methods such as continuous wavelet transform (CWT), and then processed by convolutional networks. This strategy improves the separability of non-stationary signals but introduces additional preprocessing steps and sensitivity to transform parameters.

(4) Spatio-temporal joint modeling: These approaches simultaneously capture temporal dynamics and spatial correlations across multiple sensors, thereby better reflecting the physical process of damage propagation along the structural topology. Early studies emphasized extracting spatio-temporal patterns from dense sensor networks to enhance detection and diagnostic capability, such as the spatiotemporal pattern network (STPN), which demonstrated effective modeling of spatio-temporal behavior on bridge monitoring data [24]. In recent years, various spatio-temporal feature fusion network architectures have been proposed to further strengthen cross-sensor and cross-scale representation learning.

Despite these advances, most existing approaches remain fundamentally correlation-driven and therefore tend to exhibit performance degradation when the data distribution shifts due to changes in operating conditions, environmental factors, or sensor configurations. This limitation highlights the need for learning mechanisms that explicitly promote robustness and stability across varying conditions.

2.3 Graph-based modeling of structural damage propagation

The effect of bridge damage on vibration responses is typically not confined to a local region. Changes in stiffness or connection states induced by damage can influence the responses of neighboring and even distant measurement points through component connectivity and force transmission mechanisms. Consequently, representing the sensor network or structural topology explicitly as a graph and performing information propagation and aggregation over this graph has become an important approach for capturing spatial mechanisms in SHM.

Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs) enable structural information fusion through neighborhood aggregation and attention-based weighting, providing general modeling tools for multi-point coupling in structural monitoring applications [25, 26]. In

recent years, a variety of GNN-based frameworks have been proposed for structural damage detection and localization, with a particular focus on improving localization accuracy and robustness by incorporating graph-based structural information, and these approaches have gradually been extended to real bridges and other complex structures [27–29].

However, graph-based modeling alone does not inherently address the problem of cross-condition distribution shift. When graph representations are learned primarily from observational data, they may still encode correlations associated with external factors such as vehicle speed, environmental noise, or data acquisition conditions. As a result, generalization across different operating conditions may remain limited, even when structural topology is explicitly incorporated into the model.

2.4 Cross-condition and cross-domain generalization

In response to variations in operating conditions, environmental changes, and data distribution shifts commonly encountered in SHM applications, two main methodological directions have been explored:

(1) Domain adaptation: This paradigm typically assumes that data from the target domain, either unlabeled or sparsely labeled, are available during training, and it seeks to improve transfer performance by aligning source and target data distributions through techniques such as adversarial learning, statistical distance minimization, or feature normalization. However, in practical engineering settings, target operating conditions are often unknown in advance or only partially observed, which limits the feasibility of domain adaptation in long-term monitoring scenarios.

(2) Domain generalization: This approach aims to learn representations that are insensitive to variations across multiple training environments, thereby enabling stable performance under previously unseen conditions. Domain generalization is more consistent with the long-term service characteristics of bridge structures, where training data may cover several typical operating conditions, but deployment inevitably encounters new combinations of vehicle speeds, load levels, or noise conditions. As a result, recent SHM research has increasingly focused on systematic evaluation and method development for cross-condition robustness. Datasets such as the Old ADA Bridge benchmark, which includes multiple vehicle speeds and controlled damage scenarios, have therefore become important platforms for assessing generalization performance [21–23, 30].

2.5 Causal invariant learning and invariant risk minimization (IRM)

Deep models based purely on correlation-driven learning often exploit spurious features that are statistically associated with labels but not causally related to the underlying damage mechanisms, such as spectral shifts induced by changes in vehicle speed or variations in sensor noise levels. When operating conditions change, these spurious correlations tend to break down, resulting in a substantial loss of predictive performance. From a causal learning perspective, if the structural damage mechanism (Damage \rightarrow Response) remains stable across different environments, the model should be encouraged to learn decision rules that are invariant across those environments.

IRM formalizes this idea by seeking feature representations

for which the same optimal classifier applies across multiple training environments, thereby enabling improved out-of-distribution generalization [31]. Subsequent studies have extended and analyzed IRM from both theoretical and empirical perspectives. For example, IRM games provide alternative formulations and theoretical insights [32], while empirical evaluations examine the conditions under which IRM succeeds or fails [33]. Other work has further highlighted potential limitations and risks associated with invariant learning, including scenarios in which the desired invariance cannot be reliably recovered [34, 35].

Taken together, these studies indicate that invariant learning is not a universal solution but rather a principled constraint that can guide representation learning toward features with greater stability across environments. In the context of bridge SHM, treating operating conditions as environmental variables and introducing invariance regularization on top of spatio-temporal and graph-based representation learning provides a natural mechanism for improving cross-condition generalization without requiring access to target-condition data.

2.6 Interpretability and localization evidence in SHM

In engineering applications, models are expected not only to achieve accurate classification performance but also to provide explanations that are meaningful for structural safety assessment and decision-making. In recent years, explainable deep learning has attracted increasing attention in SHM, including techniques such as attention-based weighting of key measurement points, visualization of impact paths through graph propagation, and anomaly explanation based on reconstruction errors or representation shifts.

Representative studies have emphasized post-deployment adaptive and long-term stable monitoring paradigms, such as mechanics-informed autoencoders, which demonstrate the feasibility of baseline learning and damage localization

without prior damage information [36]. When interpretability analysis is aligned with physical mechanisms, for example, damage propagation paths along the structure, it can substantially enhance the engineering credibility and practical usefulness of data-driven methods. However, under cross-condition scenarios, explanation patterns may still drift as a result of spurious correlations induced by changing operating conditions.

Therefore, jointly considering the stability of interpretability and cross-condition robustness has become an important direction for improving the reliability and usability of SHM systems in practical engineering environments.

3. METHOD

To address performance degradation in bridge damage identification caused by variations in operating conditions, a CISRL framework is developed. The framework takes multi-sensor vibration time series as input, incorporates bridge structural topology and damage propagation mechanisms, explicitly models the spatial diffusion of damage effects within the structure, and introduces cross-condition invariance constraints to encourage the learning of stable feature representations that are insensitive to operating condition changes while remaining discriminative for structural damage.

The overall framework consists of three interconnected modules: (1) a spatio-temporal feature extraction module for learning temporal and cross-sensor representations from multi-channel vibration signals; (2) a damage propagation graph modeling module for capturing the spatial diffusion and structural coupling of damage effects; and (3) a causal invariance constraint and classification module for suppressing condition-dependent spurious correlations and improving cross-condition generalization. An overview of the framework is shown in Figure 1.

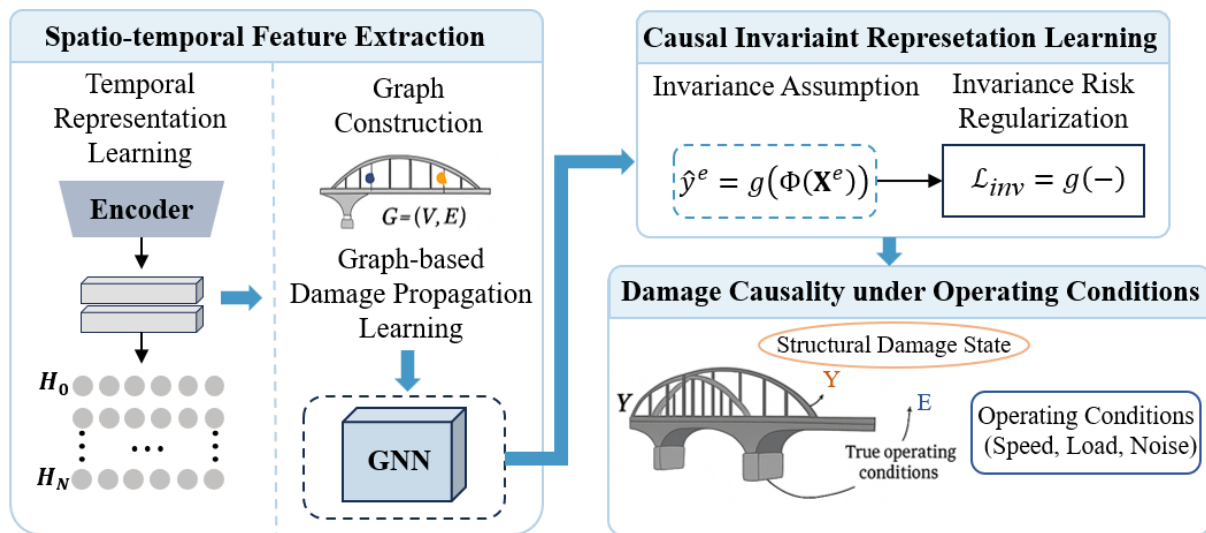


Figure 1. Method workflow

3.1 Problem definition and causal perspective

Consider the scenario where vibration responses are collected from a bridge structure under different operational conditions. Let the bridge have N accelerometers installed at key locations, and multi-channel vibration signals are

collected under operational conditions $e \in \mathcal{E}$ (such as different vehicle speeds, load levels, and environmental noise conditions):

$$\mathbf{X}^e = [x_1^e, x_2^e, \dots, x_N^e] \in \mathbb{R}^{N \times T}$$

where T represents the time sampling length, and $x_i^e \in \mathbb{R}^T$ is the vibration time series of the i -th sensor under condition e .

The true structural state of the bridge is represented by the random variable Y (such as healthy/damaged state or specific damage location), and the condition variable E represents external operational conditions. From a causal modeling perspective, structural damage is the root cause of changes in vibration response, while condition factors, as external disturbances, affect the amplitude, spectrum, and statistical distribution of the observed signals, but do not change the "damage-response" physical causal mechanism. This causal relationship can be briefly represented as:

$$Y \rightarrow X, E \rightarrow X$$

Thus, the core objective of this paper is to learn a prediction function $f(\cdot)$ under multiple operational conditions, such that its decision relies primarily on the causal influence of structural damage Y , while minimizing reliance on the condition variable E , thereby achieving stable and generalizable damage identification and localization.

3.2 Spatio-temporal feature extraction module

Bridge vibration signals usually exhibit significant non-stationarity, multi-scale features, and complex temporal dependencies. Relying solely on manually designed time-domain or frequency-domain features cannot fully represent damage information. Therefore, this paper adopts a deep time encoding network to perform end-to-end feature learning for each sensor's vibration signal.

For the i -th sensor, its temporal feature representation is:

$$h_i^e = F_t(x_i^e), h_i^e \in \mathbb{R}^d$$

where $F_t(\cdot)$ represents the time encoder, consisting of multiple layers of 1D convolution networks and attention mechanisms, used to simultaneously capture local vibration patterns and long-term dependencies. The convolution structure effectively extracts dynamic features at different time scales, while the attention mechanism further enhances the model's ability to focus on key moments and anomalous responses.

By independently encoding all sensor signals, the initial node feature matrix is obtained:

$$H_0^e = [h_1^e, h_2^e, \dots, h_N^e] \in \mathbb{R}^{N \times d}$$

This representation retains the local dynamic response information of each measurement point, providing the foundation for subsequent spatial modeling.

3.3 Damage causal propagation graph modeling

3.3.1 Graph construction based on structural topology

In real bridge structures, the impact of damage often propagates step by step along the component connection paths, and the dynamic response between different measurement points shows significant spatial correlation. To model this physical process, this paper introduces a damage causal propagation graph ($G = (V, E)$), where, the node set $V = \{v_1, v_2, \dots, v_N\}$ corresponds to sensor locations; the edge set E represents the component connection relationships and potential damage propagation paths in the bridge structure. The adjacency matrix $A \in \mathbb{R}^{N \times N}$ can be constructed based on

the bridge structural topology, component connection information, or spatial distance between measurement points, thereby explicitly injecting engineering prior knowledge into the model.

3.3.2 Damage propagation learning based on GNN

On the damage propagation graph, a GNN is used to iteratively update node features to simulate the spatial diffusion of damage within the structure. The update rule for the l -th layer of graph convolution is defined as:

$$H^{(l+1)} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right)$$

where $\tilde{A} = A + I$, \tilde{D} is the corresponding degree matrix, $W^{(l)}$ is the learnable parameter, and $\sigma(\cdot)$ is the nonlinear activation function.

Through multiple layers of graph propagation, the model can progressively aggregate information from neighboring measurement points, allowing the node features to encode both local dynamic response characteristics and structural-level damage propagation patterns, thus improving the spatial consistency and physical interpretability of damage localization.

3.4 Causal invariant representation learning

3.4.1 Engineering and causal motivation for cross-condition invariance assumption

In the practical application of SHM, vibration responses are not only influenced by the structural state but are also significantly disturbed by operational conditions such as vehicle speed, load, traffic density, and environmental noise. Vibration signals collected under different conditions often exhibit significant differences in amplitude distribution, spectral energy, and statistical characteristics, which makes models based on correlation learning prone to misclassifying condition-related features as damage features, leading to significant performance degradation under unseen conditions.

However, from the perspective of structural dynamics and engineering mechanisms, structural damage is the root cause of changes in dynamic response patterns. Although variations in operational conditions may affect the observed form of signals, they do not alter the intrinsic physical mechanism of "damage \rightarrow response." Therefore, it is reasonable to assume that:

The discriminative response patterns caused by structural damage should maintain causal consistency under different operational conditions.

Based on this understanding, this paper introduces the cross-condition invariance assumption, treating different operational conditions as different "environments" and requiring the model to learn a consistent decision-making mechanism across multiple environments. This idea is highly consistent with the invariance principle in causal inference, which states that maintaining stable causal relationships across different environments is key to achieving reliable generalization.

3.4.2 Modeling causal invariant representations

In the CISRL framework proposed in this paper, the feature extraction and damage propagation modeling modules jointly form the feature mapping function $\Phi(\cdot)$, which is used to extract high-level representations from multi-channel

vibration signals. Then, the classifier $g(\cdot)$ is used to predict the structural state:

$$\hat{y}^e = g(\Phi(X^e))$$

If the model is trained solely by minimizing classification error, the classifier may rely on different discriminative cues under different conditions, thus forming multiple "condition-specific" decision rules. To avoid this issue, this paper further introduces a causal invariance constraint that explicitly ensures the decision behavior of the model remains consistent across different conditions.

The core idea of this constraint is that, across all operational conditions, the optimal classifier should have the same discriminative direction and decision boundary, forcing the model to discard reliance on condition-specific statistical features and retain only causal features directly related to structural damage.

3.4.3 Invariant risk regularization

To achieve the above goal, this paper adopts the concept of IRM and introduces a cross-condition consistency regularization term. Specifically, the regularization term constrains the gradient of the classification loss with respect to the classifier parameters to remain consistent across different conditions:

$$\mathcal{L}_{\text{inv}} = \sum_{e \in \mathcal{E}} \|\nabla_{\theta} \mathcal{L}^e\|^2$$

where \mathcal{L}^e represents the classification loss under condition e .

This regularization term constrains the update direction of the classifier from an optimization perspective, guiding the model to converge toward a shared optimal solution across different conditions. Intuitively, this mechanism encourages the model to use only those features that have stable discriminative power across all conditions, effectively suppressing spurious correlation patterns introduced by factors such as vehicle speed changes and noise level variations.

3.5 Overall optimization objective and training strategy

Considering both structural state discrimination performance and cross-condition generalization ability, this paper incorporates both classification loss and causal invariance constraint into a unified optimization objective function:

$$\mathcal{L}_{\text{total}} = \sum_{e \in \mathcal{E}} \mathcal{L}_{\text{cls}}^e + \lambda \mathcal{L}_{\text{inv}}$$

where $\mathcal{L}_{\text{cls}}^e$ is the standard classification loss under condition e , and λ is the weight coefficient used to balance the discriminative performance and the strength of the invariance constraint.

In the actual training process, the model first learns the basic damage discrimination ability by minimizing the classification loss. Then, the invariance constraint gradually guides the model to adjust its feature representations to maintain consistent discriminative structures across different conditions. This joint optimization strategy can avoid sacrificing discriminative performance in pursuit of invariance while ensuring that the model still maintains stable predictive ability under unseen conditions.

4. RESULTS AND ANALYSIS

4.1 Damage detection results on the old ADA bridge dataset

Table 1 summarizes the quantitative performance of different methods on the Old ADA Bridge dataset for bridge damage detection, and Figure 2 visually presents the accuracy comparison among these methods. As seen, the ResNet34-1D method based on a one-dimensional convolutional network, which uses only time-domain information for feature learning, has relatively limited detection performance, with accuracy and F1-score values of 0.734 and 0.733, respectively. The CNN-LSTM model improves performance by introducing temporal dependency modeling but still struggles to adequately capture response changes caused by structural damage under complex conditions.

Table 1. Performance comparison of different methods on the old ADA bridge damage detection task

Method	Accuracy	Precision	Recall	F1-score
ResNet34-1D	0.734	0.768	0.734	0.733
CNN-LSTM	0.842	0.866	0.842	0.845
GoogLeNet (CWT)	0.826	0.834	0.826	0.824
CNN-LSTM-GoogLeNet	0.884	0.892	0.884	0.885
STTFNet	0.950	0.953	0.950	0.951
Proposed CISRL (ours)	0.964	0.967	0.964	0.965

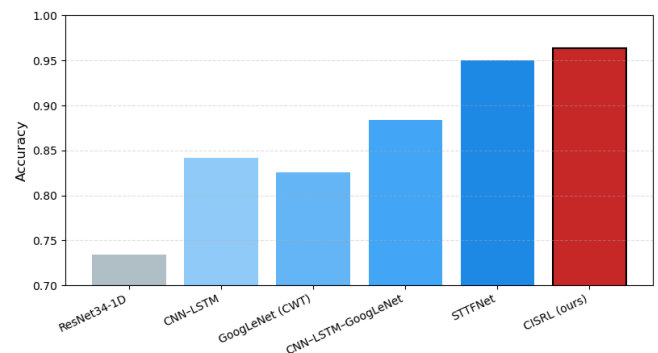


Figure 2. Damage detection results on the old ADA bridge dataset

The GoogLeNet (CWT) method, combined with time-frequency transformations, and the multi-network fusion model CNN-LSTM-GoogLeNet further enhance the detection accuracy, demonstrating the positive effect of multi-scale features on damage recognition. STTFNet, by explicitly modeling the spatiotemporal characteristics of the vibration signals, achieves a significant performance improvement on this dataset, with accuracy and F1-score values of 0.950 and 0.951, respectively.

On this basis, the CISRL method proposed in this paper achieves the best results in all evaluation metrics, with an accuracy of 0.964 and an F1-score of 0.965. Compared to STTFNet, the overall performance is improved by about 1.4%. It should be noted that this performance improvement is not due to a more complex network structure or deeper model layers, but rather the effectiveness of the causal invariance

learning mechanism in suppressing condition-related spurious features during the feature learning phase, allowing the model to focus more on the stable response patterns driven by structural damage.

4.2 Cross-condition generalization performance analysis

To further validate the model's generalization ability when the operational conditions change, cross-condition damage detection experiments were conducted. Specifically, the model was trained on data from vehicle speeds of 30 km/h and 40 km/h and tested on data from a completely unseen condition of 50 km/h. Table 2 presents a performance comparison of different methods under this cross-condition setting.

Table 2. Cross-condition damage detection performance comparison

Method	Accuracy	F1-score	Performance Drop
CNN-LSTM	0.711	0.706	-13.6%
STTFNet	0.892	0.889	-6.1%
CISRL (ours)	0.936	0.934	-3.0%

It can be observed that when there is a significant distribution shift between the training and testing conditions, traditional deep models experience a substantial performance

drop. For example, the accuracy and F1-score of the CNN-LSTM model decrease to 0.711 and 0.706, respectively, with a performance degradation of 13.6%. STTFNet mitigates this issue to some extent, but its performance still drops by about 6.1%, indicating that relying solely on spatiotemporal feature modeling cannot completely eliminate the interference caused by changes in conditions.

In contrast, the proposed CISRL method shows the least performance degradation in cross-condition testing, with accuracy and F1-score remaining at 0.936 and 0.934, respectively, and a performance drop of only 3.0%. Compared to STTFNet, the performance degradation is reduced by about 50%. This result fully verifies the effectiveness of explicitly incorporating the causal invariance constraint in improving cross-condition generalization ability, showing that the model can learn stable discriminative features that are insensitive to condition changes but highly related to structural damage.

As shown in Figure 3, this paper further analyzes the cross-condition generalization ability of different methods from the perspective of performance and the internal attention behavior of the model. Figure 3(a) presents the F1-score variation trend of different methods under unseen vehicle speed conditions. It can be observed that as the vehicle speed gradually deviates from the training conditions, the performance of both CNN-LSTM and STTFNet decreases to varying degrees, while the CISRL method maintains a more stable degradation curve.

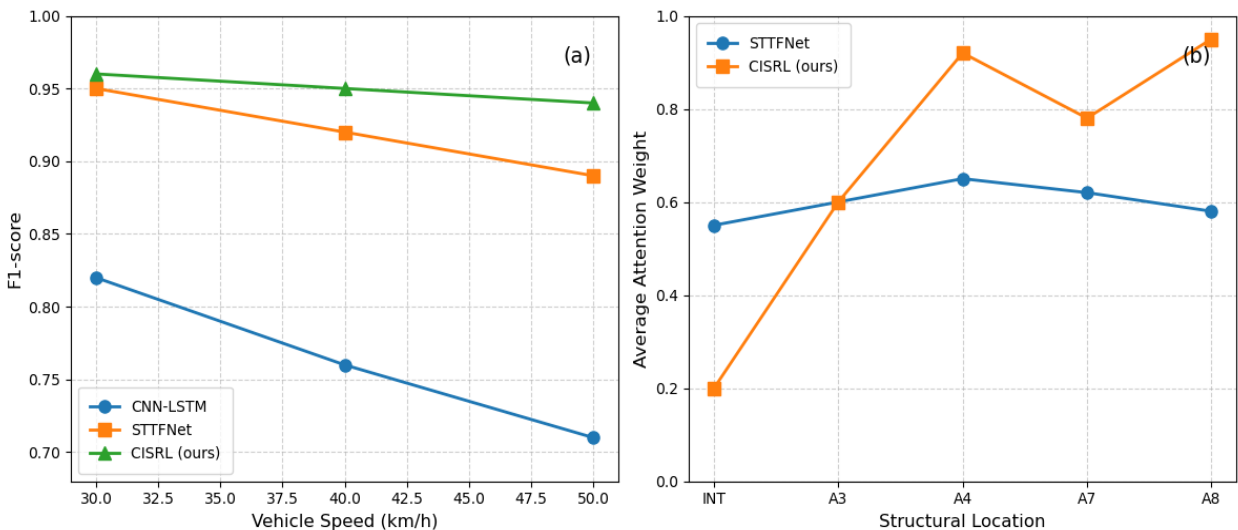


Figure 3. (a) Cross-condition performance degradation curve, (b) Attention-position curve

To further explain this phenomenon, Figure 3(b) shows the average attention weight distribution of different methods across structural positions. It can be observed that CISRL allocates significantly higher attention weights to the main damage locations (A4, A7, and A8) while significantly suppressing the response of the intact state (INT). In contrast, STTFNet's attention distribution is more dispersed. This result indicates that the causal invariance constraint helps the model continuously focus on key locations directly related to structural damage under cross-condition scenarios, effectively supporting its superior cross-condition generalization performance.

4.3 Damage localization results analysis

In addition to the overall damage detection performance,

this paper further evaluates the performance of the proposed CISRL method in damage localization tasks. Table 3 summarizes the model's overall performance metrics for damage localization, with accuracy, precision, recall, and F1-score reaching 0.931, 0.942, 0.938, and 0.940, respectively. Compared with existing representative methods based on correlation modeling (such as STTFNet, which achieves performance in the range of 0.916-0.930 for similar tasks), CISRL demonstrates stable improvements across all metrics, indicating that it improves localization accuracy while maintaining good generalization stability.

At the position level, CISRL achieves 100% localization accuracy at damage positions with more significant dynamic response features (such as A4 and A8), showing its high sensitivity to key structural damage features. For positions with relatively complex response patterns influenced by

multiple operating conditions (such as A3 and A7), the model still achieves localization accuracies of 92% and 95%, respectively, maintaining a high level of overall localization performance.

Table 3. Overall performance in damage localization task

Metric	Value
Accuracy	0.931
Precision	0.942
Recall	0.938
F1-score	0.940

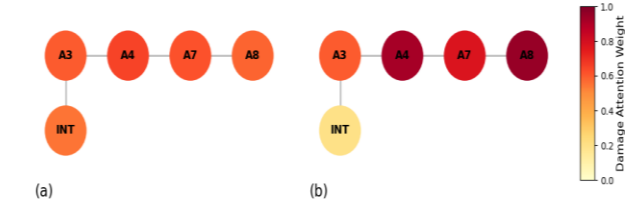


Figure 4. (a) STTFNet (Correlation-based), (b) Proposed CISRL (Causal-invariant)

In contrast, recognizing the intact structural state (INT) is relatively challenging, with an accuracy of approximately 85%. This phenomenon mainly arises from the high similarity in dynamic response characteristics between the intact state and minor damage, causing some overlap in the feature space. Even with the introduction of causal invariance constraints, fully distinguishing between these states remains challenging. This result aligns with the physical understanding in actual engineering SHM. Combining the visualization results in Figure 4, it can be further observed that, compared to correlation-driven methods, CISRL shows a more concentrated and physically consistent attention distribution across spatial dimensions, effectively suppressing redundant responses from the intact structural region. This result provides an intuitive explanation for CISRL's performance advantage in the damage localization task from the perspective of spatial propagation.

4.4 Ablation experiment analysis

To analyze the contribution of each module to the overall performance, systematic ablation experiments were conducted, and the results are shown in Table 4. When only temporal feature modeling and graph structure modeling are used without introducing causal invariance constraints, the model achieves an accuracy of 0.949 and F1-score of 0.949, indicating the significant role of spatiotemporal modeling in damage detection. However, when causal invariance constraints are introduced but graph structure modeling is removed, the model performance drops significantly, indicating that spatial structural information is crucial for damage propagation modeling and localization tasks.

Further analysis shows that causal invariance constraints play a key role in improving cross-condition performance, while graph structure modeling significantly enhances the model's ability to characterize the spatial distribution of damage. When all three modules are enabled, the model achieves the best balance between discriminative performance and generalization ability, validating the complementarity of the modules in the CISRL framework.

Table 4. Ablation experiment results for modules

Temporal	Graph	Invariance	Accuracy	F1-score
✓	✓	✗	0.949	0.949
✓	✗	✓	0.917	0.919
✗	✓	✓	0.904	0.907
✓	✓	✓	0.964	0.965

4.5 Robustness analysis

Table 5 presents the F1-score comparison between CISRL and STTFNet under different signal-to-noise ratio (SNR) conditions. As the SNR decreases, the performance of all methods declines to varying degrees. However, compared to STTFNet, CISRL maintains higher detection performance under low SNR conditions. For example, at an SNR of 5 dB, CISRL's F1-score is 0.918, while STTFNet's is only 0.884.

This result indicates that the causal invariant features learned by CISRL are more robust to noise disturbances, further verifying its applicability in complex engineering environments.

Table 5. F1-score at different noise levels

SNR (dB)	STTFNet	CISRL
20	0.948	0.960
10	0.921	0.945
5	0.884	0.918

5. DISCUSSION AND CONCLUSION

To address the degradation of generalization performance caused by variations in operating conditions in bridge damage identification, a CISRL framework has been developed. Experimental results under cross-condition testing scenarios indicate that the proposed approach exhibits smaller performance degradation than conventional correlation-driven deep models, supporting the effectiveness of causal invariance constraints in suppressing condition-related spurious features.

In addition, explicitly incorporating bridge structural topology into damage propagation graph modeling enables a more accurate characterization of the spatial diffusion patterns of damage effects, which contributes to improved damage localization performance. Ablation studies and noise robustness analyses further show that spatio-temporal feature modeling, graph-based structural propagation, and causal invariance constraints play complementary roles within the overall framework. Among these components, causal invariant learning is particularly important for enhancing cross-condition generalization, whereas graph modeling mechanisms primarily contribute to spatial discriminative capability.

Overall, the CISRL framework improves the stability and practical applicability of data-driven damage identification models under complex and varying operating conditions, while maintaining consistency with the underlying physical mechanisms of structural response. These characteristics suggest that causal-invariant learning provides a viable basis for robust and interpretable bridge SHM in real engineering environments.

ACKNOWLEDGMENT

This research was supported by multiple funding sources,

including the project “Construction of High-level Discipline Team for Environmental Safety and Governance” from the School of Management Science and Engineering, Guangxi University of Finance and Economics; the Research Basic Capacity Enhancement Program for Young and Middle-aged Teachers in Guangxi Universities (Grant No.: 2025KY0648); and the Excellent Training Program (Project No.: GPKY202402 and GPKY202411).

REFERENCES

- [1] Farrar, C.R., Worden, K. (2007). An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851): 303-315. <https://doi.org/10.1098/rsta.2006.1928>
- [2] Farrar, C., Czarnecki, J.J., Sohn, H., Hemez, F. (2002). A review of structural health monitoring literature 1996-2001. <https://api.semanticscholar.org/CorpusID:65173467>.
- [3] Doebling, S.W., Farrar, C.R., Prime, M.B., Shevitz, D.W. (1996). Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review. LA-13070-MS. <https://doi.org/10.2172/249299>
- [4] Salawu, O.S. (1997). Detection of structural damage through changes in frequency: A review. *Engineering Structures*, 19(9): 718-723. [https://doi.org/10.1016/S0141-0296\(96\)00149-6](https://doi.org/10.1016/S0141-0296(96)00149-6)
- [5] Wang, Y.Y., Qu, Y.L., Jiao, Y.B. (2024). A novel image processing-based method for wind-induced vibration response prediction and structural monitoring of long-span bridges. *Traitement du Signal*, 41(6): 3313-3326. <https://doi.org/10.18280/ts.410647>
- [6] Stubbs, N., Kim, J.T., Farrar, C.R. (1995). Field verification of a nondestructive damage localization and severity estimation algorithm. In *Proceedings-SPIE the International Society for Optical Engineering*, pp. 210-210.
- [7] Raghavan, A.C., Cesnik, C.E.S. (2007). Review of guided-wave structural health monitoring. *The Shock and Vibration Digest*, 39(2): 91-114. <https://doi.org/10.1177/0583102406075428>
- [8] Yan, Y.J., Cheng, L., Wu, Z.Y., Yam, L.H. (2007). Development in vibration-based structural damage detection technique. *Mechanical Systems and Signal Processing*, 21(5): 2198-2211. <https://doi.org/10.1016/j.ymssp.2006.10.002>
- [9] Worden, K., Manson, G. (2007). The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851): 515-537. <https://doi.org/10.1098/rsta.2006.1938>
- [10] Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., Inman, D.J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *Journal of Sound and Vibration*, 388: 154-170. <https://doi.org/10.1016/j.jsv.2016.10.043>
- [11] Cui, J., Lv, C., Du, J. (2025). Real-time structural health monitoring of steel structures using acoustic emission signals and a KAN-LSTM deep learning framework. *Engineering Structures*, 344: 121328. <https://doi.org/10.1016/j.engstruct.2025.121328>
- [12] Qin, H.W., Zou, J., He, B.G., Fu, Y., Wang, L.Z. (2023). Damage identification method of wind turbine generator system blades based on image processing technology. *Traitement du Signal*, 40(2): 825-833. <https://doi.org/10.18280/ts.400245>
- [13] Wang, X., Chen, Z., Sun, W.J., Shao, N., You, Z.Y., Xu, J.W., Yan, R.Q. (2024). A small sample piezoelectric impedance-based structural damage identification using signal reshaping-based enhance attention transformer. *Mechanical Systems and Signal Processing*, 208: 111067. <https://doi.org/10.1016/j.ymssp.2023.111067>
- [14] Xu, X.B., Pan, T.B., Zheng, Y.L., Lan, X., Zhou, Y.J., Hou, C.Y. (2025). A novel structural damage identification method based on multi-sensor data fusion and multimodal neural networks. *Engineering Structures*, 345: 121512. <https://doi.org/10.1016/j.engstruct.2025.121512>
- [15] Geirhos, R., Jacobsen, J.H., Michaelis, C., Zemel, R., Brendel, W., Bethge, M., Wichmann, F.A. (2020). Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11): 665-673. <https://doi.org/10.1038/s42256-020-00257-z>
- [16] Banad, Y.M., Sharif, S.S., Rezaei, Z. (2025). Artificial intelligence and machine learning for smart grids: From foundational paradigms to emerging technologies with digital twin and large language model-driven intelligence. *Energy Conversion and Management*: X, 28: 101329. <https://doi.org/10.1016/j.ecmx.2025.101329>
- [17] Li, X.M., Wang, Y.Y., Xing, J.D., Wang, Y.X. (2026). Causal graph inference with adaptive dynamic structure learning for mechanism-oriented fault diagnosis in dynamic industrial systems. *Reliability Engineering & System Safety*, 266(Part B): 111865. <https://doi.org/10.1016/j.res.2025.111865>
- [18] Pearl, J. (2009). *Causality*. Cambridge University Press.
- [19] Peters, J., Bühlmann, P., Meinshausen, N. (2016). Causal inference by using invariant prediction: Identification and confidence intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5): 947-1012. <https://doi.org/10.1111/rssb.12167>
- [20] Arjovsky, M., Bottou, L., Gulrajani, I., Lopez-Paz, D. (2019). Invariant risk minimization. *arXiv preprint arXiv:1907.02893*. <https://doi.org/10.48550/arXiv.1907.02893>
- [21] Kim, C.W., Zhang, F.L., Chang, K.C., McGetrick, P.J., Goi, Y. (2021). Ambient and vehicle-induced vibration data of a steel truss bridge subject to artificial damage. *Journal of Bridge Engineering*, 26(7): 04721002. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0001730](https://doi.org/10.1061/(ASCE)BE.1943-5592.0001730)
- [22] Kim, C.W., Zhang, F., Chang, K.C., McGetrick, P., Goi, Y. (2021). *Old ADA Bridge-damage_vibration_data*. Mendeley Data. <https://doi.org/10.17632/sc8whx4pvm.2>
- [23] Sohn, H., Farrar, C.R. (2001). Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 10(3): 446. <https://doi.org/10.1088/0964-1726/10/3/304>
- [24] Liu, C., Gong, Y., Laflamme, S., Phares, B., Sarkar, S. (2016). Bridge damage detection using spatiotemporal patterns extracted from dense sensor network. *Measurement Science and Technology*, 28(1): 014011. <https://doi.org/10.1088/1361-6501/28/1/014011>
- [25] Kipf, T.N., Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv*

- preprint arXiv:1609.02907.
<https://doi.org/10.48550/arXiv.1609.02907>
- [26] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y. (2017). Graph attention networks. arXiv preprint arXiv:1710.10903. <https://doi.org/10.48550/arXiv.1710.10903>
- [27] Xu, W.L., Yang, B., Wang, S.L., Bi, F.Y., Wang, S.B., Liu, L., Huang, Y.H. (2025). Multi-region transfer graph convolutional network for the localization and evaluation of multi-damage in large-scale composite structures. *Mechanical Systems and Signal Processing*, 241: 113489. <https://doi.org/10.1016/j.ymssp.2025.113489>
- [28] Del Priore, E., Lampani, L. (2025). Real-time damage detection and localization on aerospace structures using graph neural networks. *Journal of Sensor and Actuator Networks*, 14(5): 89. <https://doi.org/10.3390/jsan14050089>
- [29] Wang, C.G., Tian, X.Y., Zhou, F.N., Karimi, H.R. (2026). Intelligent fault diagnosis of bearings based on unsupervised domain adaptive adversarial graph neural network under variable operating conditions. *Measurement*, 259(Part B): 119697. <https://doi.org/10.1016/j.measurement.2025.119697>
- [30] Tomassini, E., Centofanti, G., Chellini, G., García-Macías, E., Lepori, L., Mannella, P., Salvatore, W., Ubertini, F. (2025). Key findings from long-term operational modal analysis of a landmark steel arch bridge in Italy. *Structures*, 82: 110436. <https://doi.org/10.1016/j.istruc.2025.110436>
- [31] Arjovsky, M., Bottou, L., Gulrajani, I., Lopez-Paz, D. (2019). Invariant risk minimization. arXiv preprint arXiv:1907.02893. <https://doi.org/10.48550/arXiv.1907.02893>
- [32] Ahuja, K., Shanmugam, K., Varshney, K., Dhurandhar, A. (2020). Invariant risk minimization games. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 145-155.
- [33] Choe, Y.J., Ham, J., Park, K. (2020). An empirical study of invariant risk minimization. arXiv preprint arXiv:2004.05007. <https://doi.org/10.48550/arXiv.2004.05007>
- [34] Rosenfeld, E., Ravikumar, P., Risteski, A. (2020). The risks of invariant risk minimization. arXiv preprint arXiv:2010.05761. <https://doi.org/10.48550/arXiv.2010.05761>
- [35] Kamath, P., Tangella, A., Sutherland, D., Srebro, N. (2021). Does invariant risk minimization capture invariance? In *International Conference on Artificial Intelligence and Statistics*, pp. 4069-4077.
- [36] Li, X., Bolandi, H., Masmoudi, M., Salem, T., Jha, A., Lajnef, N., Boddeti, V.N. (2024). Mechanics-informed autoencoder enables automated detection and localization of unforeseen structural damage. *Nature Communications*, 15(1): 9229. <https://doi.org/10.1038/s41467-024-52501-4>